ANOTHER ESTIMATE OF EXPECTED CRITICAL PATH LENGTH IN PERT NETWORKS

John Lindsey, II

March 1970

P-4319

ANOTHER ESTIMATE OF EXPECTED CRITICAL PATH LENGTH IN PERT NETWORKS

John Lindsey, II*
Northern Illinois University
DeKalb, Illinois

*Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of The RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The RAND Corporation as a courtesy to members of its staff.

Professor Lindsey is a consultant to the Mathematics Department of The RAND Corporation.

SUMMARY

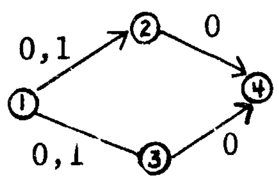
The Program Evaluation and Review Technique is a very well-known and widely-applied method for estimating the amount of time required to complete a complex project. However, if the time required to complete each subtask is a random variable, the computation of the expected time to complete the total project may be infeasible (for reasons of dimensionality). The simplest way to estimate the expected time to complete the project is to assume each subtask takes exactly its expected time to complete; and then find a critical path. This procedure, except in very simple cases, gives an underestimate of the true expected time required. In 1962 D. R. Fulkerson published a technique for estimating the expected time required which gives values intermediate between the simplest extimate and the true value. This Paper suggests another technique which, while more complex than Fulkerson's method, is more accurate in that it gives estimates which fall between Fulkerson's estimates and the true expected value.

ANOTHER ESTIMATE OF EXPECTED CRITICAL PATH LENGTH IN PERT NETWORKS

As in [1], we assume that we have a PERT network with nonnegative work times assigned along the various arcs depending on a discrete random variable u, where the probability of the value u occurring is $\psi(u)$. Number the nodes so that $j < i$ if there is an arc from node j to node i. We also assume that work times on arcs coming into any node i are independent from work times on arcs coming into a different node j. In other words, u may be considered to be a Cartesian product of random variables u_i on which work times on arcs coming into the node i depend, with $\psi(u)$ being a product of functions $\psi_i(u_i)$.

We wish to approximate the expected longest path. The simplest approach is to simply assign each arc its expected work time and calculate the longest path P_0 and its work time for this assignment. As the expectation of the sum of random variables is the sum of the expectations of the random variables, this equals the expected length of a critical (longest) path [3] if and only if for no value of u is some path longer than the path P_0. The method employed in [1] has greater flexibility in that it does not force us to use the same path for all events. In fact, the next to last node of the path is allowed to depend on the work lengths which u assigns to arcs to the final node. However, for each u, the next to

last node of a critical path also depends on the work lengths which u assigns to earlier arcs. This dependence is the reason that Fulkerson's estimate [1] may fall short of the actual expected length of a critical path. The following example illustrates this.



The arcs from 1 are allowed to take on the values 0, 1 independently with probability $\frac{1}{2}$. As there is only one assignment of work lengths along arcs to node 4, Fulkerson's method forces us to go through node 2 all the time, or through node 3 all the time. This results in the estimate $\frac{1}{2}$ while the actual expected critical length is $\frac{3}{4}$. The problem occurs when for a fixed assignment of work times on arcs to the final node and variable assignment of work times to earlier arcs, there is a close decision as to which node, j or k, should be the next to last node. In general, andom fluctuation will make a path through j longer than a path through k for some u and vice versa for other u. However, if a critical path to j tends to fluctuate the same way as a critical path to k (a critical path to j is long when a critical path to k is long) then still, little is lost in always choosing a path through u for a fixed assignment of work times along arcs to i. In summary, the Fulkerson

estimate is weakest when there is a large variance in critical path lengths to these nodes i and k, and there is little or negative correlation between these critical path lengths. It is in precisely this situation that the following modification aims to improve Fulkerson's estimate.

Let $x_j(u)$ be the length of a critical path from node 1 to node j. We define the following quantities by recursion on i or max (i,j): k_i, h_i, n_i, N_i, M_{ij} for $i \neq j$, and C_{ij} for $i > j$. Define $k_1 = h_1 = n_1 = N_1 = C_{ij} = M_{jj} = 0$. Fix i and let S be the set of nodes from which there are arcs to i. Let work times along arcs to i depend on the random variable t and work times along arcs to nodes j, $j < i$, depend on the random variable s where s and t are independent, the probability of s occurring being $\varphi(s)$, and the probability of t being $\rho(t)$. If S has one element, let $a(t) = b(t)$ be this element. Otherwise, define $a(t), b(t) \in S$ so that for $j \in S - \{a(t), b(t)\}$, $h_{a(t)} + t_{a(t)}(t) \geq h_{b(t)} + t_{b(t)}(t) \geq h_j + t_j(t)$ where $t_j(t)$ is the work time along the arc from j to i for the random variable t. In the following, if a term has 0 in its denominator, set the term equal to 0. Let $k_i = \sum_t \rho(t)\ (h_{a(t)} + t_{a(t)}(t))$, $h_i = \sum_t \rho(t) \max\ [h_{a(t)} + t_{a(t)}(t), \frac{1}{2}(h_{a(t)} + t_{a(t)}(t) + h_{b(t)} + t_{b(t)}(t)) + K(t)]$ where $K(t) = [C_{b(t)b(t)} - 2C_{a(t)b(t)} + C_{a(t)a(t)} + (h_{a(t)} + t_{a(t)}(t) - h_{b(t)} - t_{b(t)}(t))^2]/2 \max\ (M_{b(t)a(t)} + t_{b(t)}(t) - t_{a(t)}(t), M_{a(t)b(t)} + t_{a(t)}(t) - t_{b(t)}(t))$,

$n_i = \min_t (n_{a(t)} + t_{a(t)}(t))$, $N_i = \max_t [\max(N_{a(t)} + t_{a(t)}(t),$ $N_{b(t)} + t_{b(t)}(t))]$, for $i > j$, $M_{ji} = \max_t (M_{ja(t)} - t_{a(t)}(t))$ and $M_{ij} = \max_t [\max (M_{a(t)j} + t_{a(t)}(t), M_{b(t)j} + t_{b(t)}(t))]$, $C_{ii} = \sum_t \rho(t)[C_{a(t)a(t)} + (h_{a(t)} + t_{a(t)}(t) - h_i)^2] - 2(h_i - n_i)$ $(h_i - k_i)$, and for $i > j$, $C_{ij} = \sum_t \rho(t) C_{a(t)j} + (h_i - k_i)(N_j - h_j)$. Define $C_{ji} = C_{ij}$ for $j < i$. We shall show by induction on i, that for each node i, there is a function $y_i(u)$ of the random variable u, satisfying the following:

I $\quad y_i(u) \leq x_i(u)$, $n_i \leq y_i(u) \leq N_i$, $y_j(u) - y_k(u) \leq M_{jk}$, $y_{a(t)}(s) + t_{a(t)}(t) \leq y_i\ (s, t) = y_i(u) \leq \max\ (y_{a(t)}(s)$ $+ t_{a(t)}(t), y_{b(t)}(s) + t_{b(t)}(t))$;

II $\quad h_i = \bar{y}_i = \sum_u \psi(u) y_i(u)$;

III $\quad \sum_u \psi(u)(y_i(u) - h_i)^2 \geq C_{ii}$; and

IV $\quad \sum_u \psi(u)(y_j(u) - h_j)(y_k(u) - h_k) \leq C_{jk}$.

Assume that the above statements are true when i, j, and k are replaced by integers less than i. For each s, t the longest path from node 1 to node i has length $x_{j(s,\ t)}(s) +$ $t_{j(s,\ t)}(t)$ for some node $j(s, t)$ with an arc from node $j(s, t)$ to node i. Then $x_{j(s,\ t)}(s) + t_{j(s,\ t)}(t) \geq x_{a(t)}(s)$ $+ t_{a(t)}(t)$. Similarly, $x_{j(s,\ t)}(t) + t_{j(s,\ t)}(t) \geq \max$ $(y_{a(t)}(s) + t_{a(t)}(t), y_{b(t)}(s) + t_{b(t)}(t)) = \frac{1}{2}(y_{a(t)}(s) +$

$t_{a(t)}(t) + y_{b(t)}(s) + t_{b(t)}(t)) + \frac{1}{2}|y_{b(t)}(s) + t_{b(t)}(t) - y_{a(t)}(s) - t_{a(t)}(t)|$. Then $h_{a(t)} + t_{a(t)}(t) = \sum_s \varphi(s)(y_{a(t)}(s) + t_{a(t)}(t)) \leq \sum_s \varphi(s) \max\, [y_{a(t)}(s) + t_{a(t)}(t),\ y_{b(t)}(s) + t_{b(t)}(t)]$

$= \frac{1}{2} \sum_s \varphi(s)[y_{a(t)}(s) + t_{a(t)}(t) + y_{b(t)}(s) + t_{b(t)}(t)] +$

$\frac{1}{2} \sum \varphi(s)|y_{a(t)}(s) + t_{a(t)}(t) - y_{b(t)}(s) - t_{b(t)}(t)|$

$= \frac{1}{2}[h_{a(t)} + t_{a(t)}(t) + h_{b(t)} + t_{b(t)}(t)] +$

$\frac{1}{2} \sum_s \varphi(s)|y_{a(t)}(s) + t_{a(t)}(t) - y_{b(t)}(s) - t_{b(t)}(t)|$.

Also, $\frac{1}{2} \sum_s \varphi(s)|y_{b(t)}(s) + t_{b(t)}(t) - y_{a(t)}(s) - t_{a(t)}(t)|$

$\geq \sum_s \varphi(s)[y_{b(t)}(s) + t_{b(t)}(t) - y_{a(t)}(s) - t_{a(t)}(t)]^2/2 \max$

$[M_{b(t)a(t)} + t_{b(t)}(t) - t_{a(t)}(t),\ M_{a(t)b(t)} + t_{a(t)}(t) - t_{b(t)}(t)]$

and $\sum_s \varphi(s)[y_{b(t)}(s) + t_{b(t)}(t) - y_{a(t)}(s) - t_{a(t)}(t)]^2$

$= \sum_s \varphi(s)\{[(y_{b(t)}(s) - h_{b(t)}) - (y_{a(t)}(s) - h_{a(t)})] + [(h_{b(t)} + t_{b(t)}(t)) - (h_{a(t)} + t_{a(t)}(t))]\}^2$

$= \sum_s \varphi(s)(y_{b(t)}(s) - h_{b(t)})^2 - 2 \sum_s \varphi(s)(y_{b(t)}(s) - h_{b(t)})(y_{a(t)}(s) - h_{a(t)}) + \sum_s \varphi(s)(y_{a(t)}(s) - h_{a(t)})^2 + [h_{b(t)} + t_{b(t)}(t) - (h_{a(t)} + t_{a(t)}(t))]^2$

$\geq C_{b(t)b(t)} - 2C_{a(t)b(t)} + C_{a(t)a(t)} + [h_{b(t)} + t_{b(t)} - (h_{a(t)} + t_{a(t)}(t)]^2$.

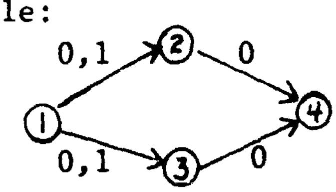
This shows that we may define $y_i(s, t)$ to satisfy the following:

$y_{a(t)}(s) + t_{a(t)}(t) \leq y_{i(s,t)} \leq \max\, (y_{a(t)}(s) + t_{a(t)}(t), y_{b(t)}(s) + t_{b(t)}(t)) \leq \max\, [x_{a(t)}(s) + t_{a(t)}(t), x_{b(t)}(s) + t_{b(t)}(t)] \leq x_i(s, t)$, and $h_i = \bar{y}_i = \sum_{s,t} \rho(t)\varphi(s)y_i(s,t) \leq \sum_{s,t} \rho(t)\varphi(s)x_i(s,t) = \bar{x}_i$. This shows by induction that h_i lies between the actual expected length of a critical path and Fulkerson's estimate for it.

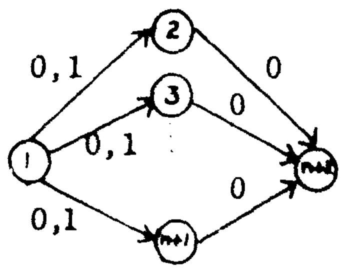
Furthermore, $n_i = \min_t [n_{a(t)} + t_{a(t)}(t)] \leq \min_{s,t} [y_{a(t)}(s) + t_{a(t)}(t)] \leq \min_{s,t} y_i(s, t) \leq \max_{s,t} y_i(s, t) \leq \max_{s,t} (\max(y_{a(t)}(s) + t_{a(t)}(t), y_{b(t)}(s) + t_{b(t)}(t)]$
$\leq \max_t [\max (N_{a(t)} + t_{a(t)}(t), N_{b(t)} + t_{b(t)}(t))] = N_i$.
For $i > j$, $M_{ji} = \max_t (M_{ja(t)} - t_{a(t)}(t))$
$\geq \max_{s,t} [y_j(s) - y_{a(t)}(s) - t_{a(t)}(t)] \geq \max_{s,t} [y_j(s) - y_i(s, t)]$
and $M_{ij} = \max_t [\max (M_{a(t)j} + t_{a(t)}(t), M_{b(t)j} + t_{b(t)}(t))]$
$\geq \max_{s,t} [\max (y_{a(t)}(s) - y_j(s) + t_{a(t)}(t), y_{b(t)}(s) - y_j(s) + t_{b(t)}(t))]$
$= \max_{s,t} [\max(y_{a(t)}(s) + t_{a(t)}(t), y_{b(t)}(s) + t_{b(t)}(t)) - y_j(s)]$
$\geq \max_{s,t} [y_i(s, t) - y_j(s)]$.
Also, $C_{ii} = \sum_t \rho(t)[C_{a(t)a(t)} + (h_{a(t)} + t_{a(t)}(t) - h_i)^2] - 2(h_i - n_i)(h_i - k_i)$
$\leq \sum_{s,t} \rho(t)\varphi(s)[y_{a(t)}(s) - h_{a(t)} + (h_{a(t)} + t_{a(t)}(t) - h_i)]^2$
$- 2(h_i - n_i)[h_i - \sum_{s,t} \rho(t)\varphi(s)(y_{a(t)}(s) + t_{a(t)}(t))]$
$= \sum_{s,t} \rho(t)\varphi(s)(y_{a(t)}(s) + t_{a(t)}(t) - h_i)^2$
$- \sum_{s,t} \rho(t)\varphi(s)(y_i(s, t) - y_{a(t)}(s) - t_{a(t)}(t))(2h_i - 2n_i)$
$\leq \sum_{s,t} \rho(t)\varphi(s)(y_{a(t)}(s) + t_{a(t)}(t) - h_i)^2$
$- \sum_{s,t} \rho(t)\varphi(s)(y_i(s, t) - y_{a(t)}(s) - t_{a(t)}(t))(2h_i - y_i(s, t) - y_{a(t)}(s) - t_{a(t)}(t))$
$= \sum_{s,t} \rho(t)\varphi(s)(y_i(s, t) - h_i)^2$. Finally, for $i > j$, C_{ij}
$= \sum_t \rho(t)C_{a(t)j} + (h_i - k_i)(N_j - h_j)$
$\geq \sum_t [\rho(t) \sum_s \varphi(s)(y_{a(t)}(s) - h_{a(t)})(y_j(s) - h_j)]$
$+ \sum_{s,t} \rho(t)\varphi(s)(y_i(s, t) - y_{a(t)}(s) - t_{a(t)}(t))(N_j - h_j)$
$\geq \sum_t [\rho(t) \sum_s \varphi(s)(y_{a(t)}(s) - h_{a(t)})(y_j(s) - h_j)]$

$$+\sum_{s,t}\rho(t)\varphi(s)(y_i(s,\ t) - y_{a(t)}(s) - t_{a(t)}(t))(y_j(s) - h_j)$$
$$+\sum_t \rho(t)(h_{a(t)} + t_{a(t)}(t) - h_i)\sum_s \varphi(s)(y_j(s) - h_j)$$
$$=\sum_{s,t}\rho(t)\varphi(s)(y_i(s,\ t) - h_i)(y_j(s) - h_j).$$

For the example:



$h_2 = k_2 = h_3 = k_3 = \frac{1}{2}$, $n_2 = n_3 = 0$, $N_2 = N_3 = 1$, $M_{12} = M_{13} = 0$, $M_{21} = M_{31} = 1$, $M_{23} = M_{32} = 1$, $C_{32} = 0$, $C_{22} = C_{33} = \frac{1}{4}$, and $h_4 = \frac{1}{2}(\frac{1}{2} + 0 + \frac{1}{2} + 0) + (\frac{1}{4} - 0 + \frac{1}{4} + (\frac{1}{2} + 0 - \frac{1}{2} - 0)^2)/2 = 3/4$, the actual expected critical length. Naturally, the example exaggerates the accuracy of our estimate. One problem is that for some assignment of work times to arcs into a node i, there may be a close decision between three or more nodes as to which should immediately precede i on a critical path. In the example



our estimate is still 3/4, but the actual expected critical path length is $1 - 2^{-n}$. Without using the assumption that work times on arcs in bundles leading into nodes j, $j < i$, are independent, it is not possible to use the correlations between critical path lengths coming in two nodes to calculate how much critical paths through c(t) to i increase

$\max\,(y_{a(t)}(s) + t_{a(t)}(t),\ y_{b(t)}(s) + t_{b(t)}(t)$; that is, by how much $\sum_s \varphi(s)\max\,(y_{a(t)}(s) + t_{a(t)}(t),\ y_{b(t)}(s) + t_{b(t)}(t),\ y_{c(t)}(s) + t_{c(t)}(t))$ exceeds $\sum_s \varphi(s)\max\,(y_{a(t)}(s) + t_{a(t)}(t),\ y_{b(t)}(s) + t_{b(t)}(t))$. In fact, suppose that $t_{a(1)}(1) = t_{b(1)}(1) = t_{c(1)}(1) = 0$ and s takes on four values with probability $\frac{1}{4}$.

s =	I	II	III	IV
$x_{a(1)}(s) = y_{a(1)}(s) =$	0	0	1	1
$x_{b(1)}(s) = y_{b(1)}(s) =$	0	1	0	1
$x_{c(1)}(s) = y_{c(1)}(s) =$	0	1	1	0

Then there is no correlation between any two of the following: $y_{a(t)}(s)$, $y_{b(1)}(s)$, and $y_{c(1)}(s)$. However, it is never necessary to use c(1) in a critical path; regardless of s, there is a critical path through a or b. However, in some sense there is correlation, as $C_{a(1)b(1)c(1)} = \sum_s \varphi(s)(y_{a(1)}(s) - \overline{y}_{a(1)})(y_{b(1)}(s) - \overline{y}_{b(1)})(y_{c(1)}(s) - \overline{y}_{c(1)}) = \frac{1}{4}(-\frac{1}{2})(-\frac{1}{2})(-\frac{1}{2}) + 3(\frac{1}{4})(\frac{1}{2})(\frac{1}{2})(-\frac{1}{2}) = -1/8$. There are two problems in trying to use this. First, we would have to calculate an expression C_{ijk} for each triple of nodes. Secondly, we would have the problem of using a third degree term C_{ijk} with the second degree terms C_{ij}.

We have already experienced these difficulties to some extent. We have to calculate expressions C_{ij} and M_{ij} for each pair of nodes, whereas, in Fulkerson's estimate, it was only necessary to calculate f_i for each node. Another

problem was that we were able to calculate $\sum_s \varphi(s)(y_{b(t)}(s) + t_{b(t)}(t) - y_{a(t)}(s) - t_{a(t)}(t))^2$, but we needed $\sum_s \varphi(s)|y_{b(t)}(s) + t_{b(t)}(t) - y_{a(t)}(s) - t_{a(t)}(t)|$ instead. Our solution was to bound the ratio by the $\max_{s,t}|y_{b(t)}(s) + t_{b(t)}(t) - y_{a(t)}(s) - t_{a(t)}(t)|$. With more simplicity and less sharpness, we could have used $\max\,(N_{b(t)} + t_{b(t)}(t) - n_{a(t)} - t_{a(t)}(t),\ N_{a(t)} + t_{a(t)}(t) - n_{b(t)} - t_{b(t)}(t))$ instead of $\max\,(M_{b(t)a(t)} + t_{b(t)}(t) - t_{a(t)}(t),\ M_{a(t)b(t)} + t_{a(t)}(t) - t_{b(t)}(t))$ as a bound for this ratio. Either of these bounds does not lose very much sharpness if $y_{b(t)}(s) + t_{b(t)}(t) - y_{a(t)}(s) - t_{a(t)}(t)$ as a function of the random variable s has a high probability of taking on a value near the endpoints of its range, as in our example. However, this will not usually be the case, particularly if we have a continuous random variable or if the node i is far from node 1 so that lots of random fluctuations make $y_{b(t)}(s) + t_{b(t)}(t) - y_{a(t)}(s) - t_{a(t)}(t)$ behave nearly like a continuous distribution. One solution is to assume that $y_{b(t)}(s) + t_{b(t)}(t) - y_{a(t)}(s) - t_{a(t)}(t)$ is distributed for fixed t as a normal distribution. If the random fluctuations add, then this is a very plausible assumption, if the node i is far from node 1, by the central-limit theorem. Now $y_{b(t)}(s) + t_{b(t)}(t) - y_{a(t)}(s) - t_{a(t)}(t)$ has mean $\bar{x} = h_{b(t)} + t_{b(t)}(t) - h_{a(t)}(t), - t_{a(t)}(t)$ and $\sigma^2 \geq C_{b(t)b(t)} - 2C_{a(t)b(t)} + C_{a(t)a(t)}$. Let $r = h_{a(t)} + t_{a(t)}(t) - h_{b(t)} - t_{b(t)}(t)$. Then $\sum_s \varphi(s)|y_{b(t)}(s) + t_{b(t)}(t) - y_{a(t)}(s) - t_{a(t)}(t)|$

approximately equals

$$\int_{-\infty}^{\infty} |y_{b(t)}(s) + t_{b(t)}(t) - y_{a(t)}(s) - t_{a(t)}(t)| \frac{1}{\sqrt{2\pi}\,\sigma}$$

$$e^{-[y_{b(t)}(s) - h_{b(t)} - y_{a(t)}(s) + h_{a(t)}]^2/2\sigma^2}$$

$$d(y_{b(t)}(s) - y_{a(t)}(s))$$

$$= \int_{r}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} (x - r)\ e^{-x^2/2\sigma^2} dx$$

$$+ \int_{-\infty}^{r} \frac{1}{\sqrt{2\pi}\,\sigma} (r - x)\ e^{-x^2/2\sigma^2} dx$$

$$= \frac{2\sigma}{\sqrt{2\pi}}\ e^{-r^2/2\sigma^2} + \frac{2r}{\sqrt{2\pi}} \int_{0}^{r/\sigma} e^{-x^2/2} dx$$

where the last term can be expressed in terms of the error function:

$$\frac{1}{\sqrt{2\pi}} \int_{0}^{r/\sigma} e^{-x^2/2} dx = \tfrac{1}{2} - \mathrm{err}(r/\sigma)$$

The expression

$$f(r,\sigma) = \frac{\sigma}{\sqrt{2\pi}} e^{-r^2/2\sigma^2} + \frac{r}{\sqrt{2\pi}} \int_o^{r/\sigma} e^{-x^2/2} dx \text{ if } \sigma^2 > 0,$$

$f(r,\sigma) = \frac{r}{2} \geq 0$ if $\sigma^2 \leq 0$, where $r = h_{a(t)} + t_{a(t)}(t) -$

$h_{b(t)} - t_{b(t)}(t)$ and $\sigma^2 = C_{b(t)b(t)} - 2C_{a(t)b(t)} + C_{a(t)a(t)}$

may be used to replace $K(t)$ in recursively defining h_i. Note that

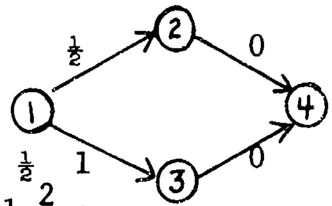
$$\frac{\partial f}{\partial \sigma} = \frac{1}{\sqrt{2\pi}} e^{-r^2/2\sigma^2} + \frac{r^2}{\sigma^2\sqrt{2\pi}} e^{-r^2/2\sigma^2} - \frac{r^2}{\sigma^2\sqrt{2\pi}} e^{-r^2/2\sigma^2}$$

$$= \frac{1}{\sqrt{2\pi}} e^{-r^2/2\sigma^2} > 0.$$

Therefore, $f(r, \sigma)$ increases as σ increases, and $f(r, \sigma) \geq \lim_{\sigma \to 0} f(r, \sigma) = \frac{r}{2} = f(r, 0)$. Therefore, we may define

$$h_i = \sum_t \rho(t)[\tfrac{1}{2}(h_{a(t)} + t_{a(t)} + h_{b(t)} + t_{b(t)}(t)) + f(r, \sigma)].$$

Also, our conservative estimates of σ give a conservative estimate for $f(r, \sigma)$. Advantages of the use of $f(r, \sigma)$ are closer estimates of the expected critical patn lengths and avoidance of the use of the expressions M_{ij} and M_{ji}. The major disadvantage is that we no longer have a guarantee that our estimate for the expected critical path length, will be optimistic, although it usually will be. In fact, consider the example



0	½	1
p	$1-\frac{2}{p}$	p

where the arc from 1 to 3 takes on the work time 0, p of the time; 1, p of the time; and ½ the rest of the time. Then $C_{22} = 0$, $C_{33} = \frac{1}{4}(2p)$. The actual expected critical path length is $\frac{1}{2} + p/2$. The new estimate for h_4 is $\frac{1}{2}(\frac{1}{2} + 0 + \frac{1}{2} + 0) + \frac{\sqrt{p/2}}{\sqrt{2\pi}}$. As p gets small, sooner or later this exceeds the expected critical path length.

Even when there are too many nodes to keep track of the C_{ij} for each pair of nodes, it is still possible to carry through this procedure for the first hundred or so nodes, label the h_i as f_i, and continue with Fulkerson's recursive definition [1] of the f_i. When our original method of defining the h_i is used, we still obtain an estimate which is no larger than the expected critical path length.

REFERENCES

1. Fulkerson, D. R., "Expected Critical Path Lengths in PERT Networks," Opns. Res. 10, 808-817 (1962)

2. MacCrimmon, K. R. and C. A. Ryavec, An Analytical Study of the PERT Assumptions, The RAND Corporation RM-3408-PR Dec. 1962.

3. Malcolm, D. G., J. H. Roseboom, C. E. Clark, and W. Frazar, "Application of a Technique for Research and Development Program Evaluation," Opns. Res. 7, 646-669 (1959)